Instituto de Ingeniería del Conocimiento



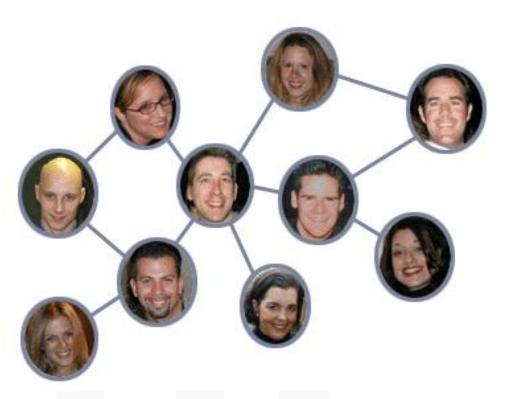


¿Qué es una red social?

Es una manera diferente de mirar a una colectividad

Colectividad = nodos + enlaces

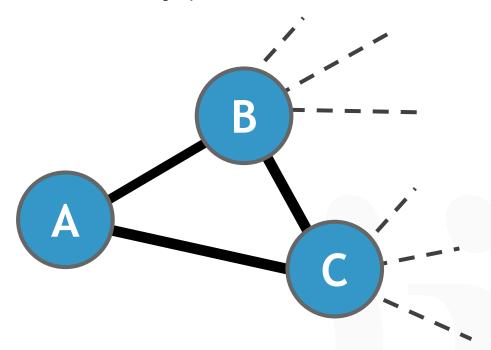
- Nodos
 - Personas
 - Organizaciones
 - ■Compañías
- Enlaces:
 - Operaciones:
 - •Llamadas/emails intercambiados
 - Transacciones
 - Relaciones, amistad
 - Afiliaciones
 - Membresía común





El valor de una red social

¿Por qué mirar a una colectividad como una red social?
 Por que el valor de una red social es mayor que la suma de sus partes.
 Es por tanto un sistema complejo



¿Cuanta de la actividad de A con B y C depende del hecho de que B y C están conectados?



Las Redes Sociales tienen una rica estructura

• "Pequeño mundo": el diámetro de una red social es pequeño en comparación con el número de nodos.

• Heterogéneas: no todos los nodos en la red tienen las mismas propiedades estructurales locales. Existe una fracción significativa de nodos altamente conectados (hubs).

Clusterizadas: a corto y medio rango, las redes sociales presentan una gran densificación debido a la existencia de comunidades, transitividad, etc.



Propiedades de una red social

- Pequeño mundo
 - Incluso en grandes redes sociales, la distancia social promedio es muy pequeña
 - Milgram 1967 (usando correo normal)
 - Distancia promedio: seis grados de separación
 - Dodds et al 2003 (usando e-mail)
 - Distancia promedio: seis grados de separación (!)

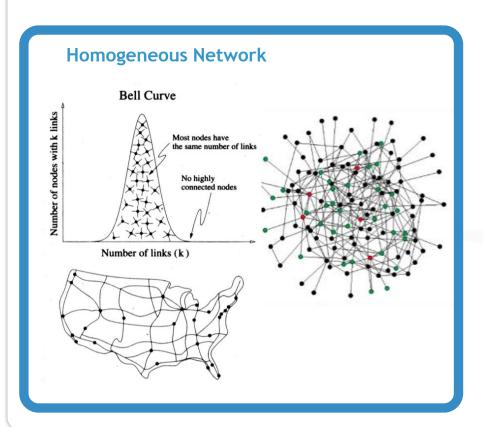


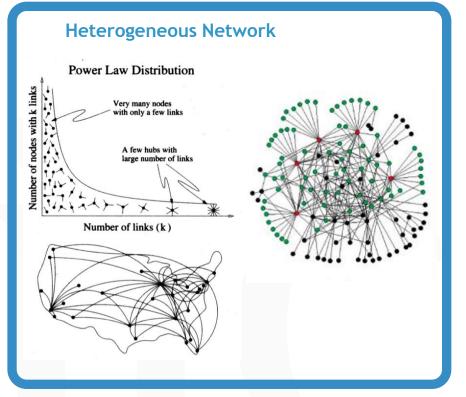




Propiedades de las redes sociales

- Las redes sociales son muy heterogéneas
 - Ley de Pareto (o ley de potencias):
 - 20% de los nodos tienen el 80% de la conectividad social
 - 80% de los nodos sólo tienen el 20% de la conectividad social

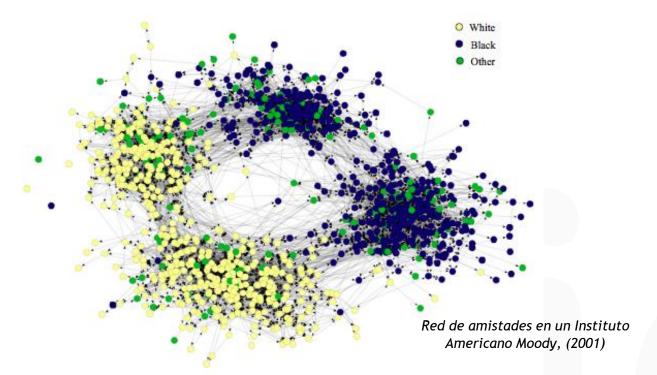






Propiedades de las redes sociales

- Existencia de Comunidades
- Grupos de nodos que tienen una diferente densidad de conexiones dentro del grupo que con el resto de la red



- Transitividad: las redes sociales son densas localmente
- La probabilidad de que dos vecinos de un nodo sean a su vez vecinos es muy alta (20%)



Análisis de redes sociales

- Representación y medición de redes sociales
 - Representación: De los datos a las relaciones y nodos
 - ¿Cuando un nodo es un nodo?
 - ¿Cuándo un enlace es un enlace?
 - Medición
 - Local
 - Conectividad social en la vecindad de un nodo
 - Clustering local o transitividad
 - Global
 - Caminos mínimos, diámetro de la red social
 - Centralidad del nodo (centros, puentes, fronteras, etc.)
 - Identificación de comunidades
 - Visualización



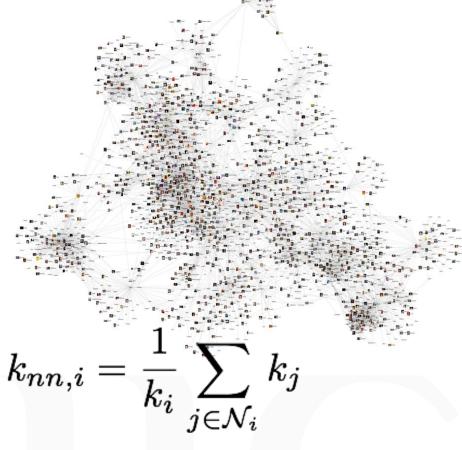
Midiendo:: Conectividad

- A primeros vecinos
 - Conectividad del nodo k_i
 - Distribución de conectividad P(k)
 - Momentos de la distribución:

$$\langle k^n \rangle = \sum_k k^n P(k)$$

- A segundos vecinos
 - Conectividad promedio de los vecinos
 - Probabilidad condicionada $P(k^\prime|k)$
 - Conectividad promedio de los vecinos condicionada

$$k_{nn}(k) = \sum_{k'} k' P(k'|k)$$



C 2011



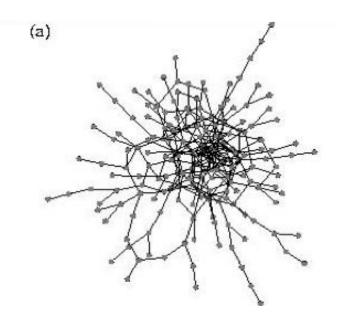
Midiendo:: Conectividad

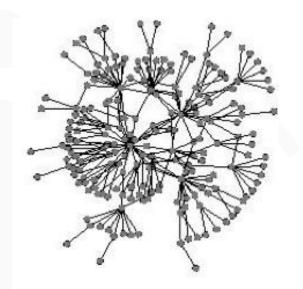
Asortatividad

- Los vecinos más próximos muestran una correlación en la conectividad social
 - Red asortativa: nodos que están muy conectados tienden a estar rodeados de nodos con alta conectividad

$$P(k'|k) = k'P(k')/\langle k \rangle$$

 Redes disortativas: nodos que están muy conectados tienden a estar rodeados de nodos con baja conectividad



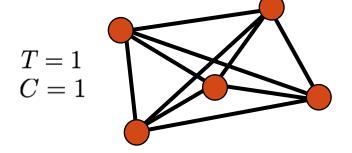




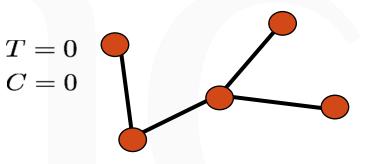
Midiendo:: Clustering

- Los vecinos de un nodo tienden a estar conectados entre ellos
 - Transitividad: probabilidad de que dos vecinos de un mismo nodo sean vecinos entre ellos

$$T = \frac{3 \times \# \text{ of triangles}}{\# \text{ of connected triples of vertices}}$$



$$T = 3 \times \frac{1}{8} = \frac{3}{8}$$



DIC 2011



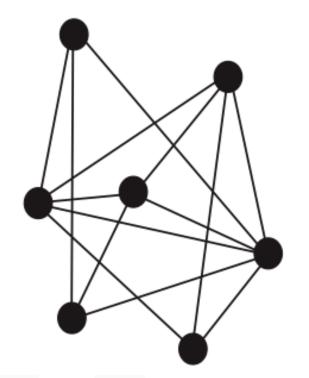
Midiendo :: Caminos mínimos, diámetro de la red

- Longitud del camino mínimo entre dos nodos $\,d_{ij}$
- Diámetro $D = \max_{i
 eq j} d_{ij}$
- Promedio de la longitud de los caminos mínimos

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$$

Eficiencia del grafo

$$E = \frac{1}{N(N-1)} \sum_{ij} \frac{1}{d_{ij}}$$



DIC 2011

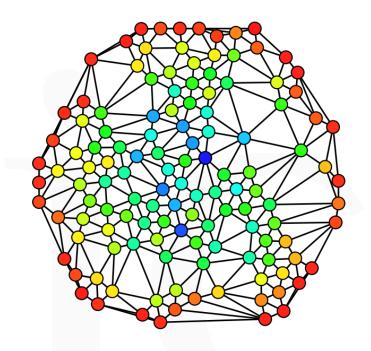


Midiendo :: Centralidad

- Identificar los "centros" de la red
- Varias maneras de hacerlo:
 - Intermediación
 - Número de caminos mínimos entre $_{j}$ y $_{k:}$ n_{jk}
 - Número de caminos mínimos entre j y k que pasan por i: $n_{jk}(i)$
 - Grado de intermediación

$$b_i = \sum_{j \neq k} \frac{n_{jk}(i)}{n_{jk}}$$

Autovectores (PageRank)





Network film actors

company directors

math coauthorship

physics coauthorship

biology coauthorship

telephone call graph

email address books

student relationships

email messages

sexual contacts

WWW nd.edu

WWW Altavista

citation network

Internet

power grid

train routes

Technological

Roget's Thesaurus

word co-occurrence

software packages

electronic circuits

metabolic network

marine food web

neural network

protein interactions

freshwater food web

peer-to-peer network

software classes

Ejemplos de métricas sobre redes sociales

Туре

undirected

undirected

undirected

undirected

undirected

undirected

undirected

undirected

directed

directed

directed

directed

directed

directed

undirected

undirected

undirected

undirected

undirected

undirected

undirected

undirected

directed

directed

directed

directed

directed

10697

4941

1439

1377

880

765

135

307

92

2115

24097

587

n	m	z	l	α	$C^{(1)}$	$C^{(2)}$	r	Ref(s).
449913	25 516482	113.43	3.48	2.3	0.20	0.78	0.208	[20, 415]
7673	55 392	14.44	4.60	_	0.59	0.88	0.276	[105, 322]
253 339	496 489	3.92	7.57	_	0.15	0.34	0.120	[107, 181]
52909	245 300	9.27	6.19	_	0.45	0.56	0.363	[310, 312]
1520251	11803064	15.53	4.92	_	0.088	0.60	0.127	[310, 312]
47 000 000	80 000 000	3.16		2.1				[8, 9]
59912	86 300	1.44	4.95	1.5/2.0		0.16		[136]
16881	57 029	3.38	5.22		0.17	0.13	0.092	[320]
573	477	1.66	16.01	_	0.005	0.001	-0.029	[45]
2810				3.2				[264, 265]
269 504	1 497 135	5.55	11.27	2.1/2.4	0.11	0.29	-0.067	[14, 34]
203 549 046	2 130 000 000	10.46	16.18	2.1/2.7				[74]
783 339	6716198	8.57		3.0/-				[350]
1022	5 103	4.99	4.87		0.13	0.15	0.157	[243]
460 902	17000000	70.13		2.7		0.44		[119, 157]

2.5

3.0

 $^{2.1}$

2.2

 $^{2.4}$

1.6/1.4

0.035

0.070

0.033

0.010

0.012

0.090

0.072

0.16

0.40

0.18

0.10

clustering

0.39

0.080

0.69

0.082

0.012

0.030

0.011

0.67

0.071

0.23

0.48

0.28

diámetro

asortatividad

-0.189

-0.003

-0.033

-0.016

-0.119

-0.154

-0.366

-0.240

-0.156

-0.263

-0.326

-0.226

[86, 148]

[415]

[365]

[317]

[394]

[155]

[213]

[211]

[203]

[271]

[415, 420]

[6, 353]

Extraído de "Structure and Function of Complex Networks" M.E.J. Newman 2003

31992

6594

19603

1723

2213

53248

1296

3686

2240

598

997

2359

5.98

2.67

66.79

1.20

1.61

4.34

1.47

9.64

2.12

4.43

10.84

7.68

3.31

2.16

2.42

1.51

4.28

2.56

6.80

2.05

1.90

3.97

11.05

18.99



¿Qué cuestiones son relevantes en el análisis de redes sociales?

Complejidad estructural

- Detección y análisis de comunidades: identificación de grupos densamente conectados.
- Capital social, centralidad: identificación de nodos que juegan papeles importantes en la estructura de la red (centros, puentes, etc.)

Complejidad dinámica

- ¿Cómo fluye la información en las redes sociales?
- ¿Cómo navega la gente en la red social buscando talento, experiencia, contactos, información, etc.?
- ¿Cómo de eficiente es la red social en la diseminación de innovaciones, ideas, opiniones, etc.?



Software de Análisis de Redes Sociales

- Listas de recursos
 - General
 - Cytoscape: http://www.cytoscape.org/cgi-bin/moin.cgi/Network analysis links
 - Center for complex network research http://www.nd.edu/~networks/resources.htm
 - Infovis: http://iv.slis.indiana.edu/sw/#packages
 - Visualización
 - VisualComplexity: http://www.visualcomplexity.com/vc/links.cfm
 - Cosin Project: http://www.cosinproject.org/
- Software genérico
 - Pajek: http://vlado.fmf.uni-lj.si/pub/networks/pajek/
 - UCINET: http://www.analytictech.com/downloaduc6.htm
 - DyNet: http://www.atalab.com/software/
- Visualización
 - Graphviz: http://www.graphviz.org/
 - Walrus (Caida):
 http://www.caida.org/tools/visualization/walrus/
 - JGraph: http://www.jgraph.com/
 - InFlow: http://www.orgnet.com



References

Introductory (general audience)

- M. Buchanan, Nexus: Small Worlds and the Groundbreaking Science of Networks, Norton, New York, 2002.
- A.-L. Barabási, Linked: The New Science of Networks, Perseus, Cambridge, 2002.
- D.J. Watts, Six Degrees: The Science of a Connected Age, Norton, New York, 2003.

Reviews on Social Networks

- [1] S.H. Strogatz, Nature 410 (2001) 268.
- [2] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47.
- [3] S.N. Dorogovtesev, J.F.F. Mendes, Adv. Phys. 51 (2002) 1079.
- [4] M.E.J. Newman, SIAM Rev. 45 (2003) 167.
- [5] D.J. Watts, Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University Press, Princeton, NJ, 1999.
- [6] S. Bornholdt, H.G. Schuster (Eds.), Handbook of Graphs and Networks: From the Genome to the Internet, Wiley-VCH, Germany, 2003.
- [7] S.N. Dorogovtesev, J.F.F. Mendes, Evolution of Networks, Oxford University Press, Oxford, 2003.
- [8] R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach, Cambridge University Press, Cambridge, 2004.
- [9] R. Pastor-Satorras, M. Rubi, A. Diaz-Guilera (Eds.), Statistical Mechanics of Complex Networks, Springer, Berlin, 2003.
- [10] E. Ben-Naim, H. Frauenfelder, Z. Toroczkai (Eds.), Complex Networks, Springer, Berlin, 2004.



Referencias

Sobre Análisis de Redes Sociales

- S. Wasserman, K. Faust, Social Networks Analysis, Cambridge University Press, Cambridge, 1994.
- J. Scott, Social Network Analysis: A Handbook, 2nd ed., Sage Publications, London, 2000.

Algoritmos de grafos

- T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, Introduction to Algorithms, MIT University Press, Cambridge, 2001.
- R. Sedgewick, Algorithms in C++. Part 5: Graph Algorithms, Addison-Wesley, Boston MA, 1988.
- R.K. Ahuja, T.L. Magnati, J.B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1993.